The Annals of Probability
1992, Vol. 20, No. 1, 137–151

UNIFORM CONVERGENCE OF MARTINGALES IN THE BRANCHING RANDOM WALK

BY J. D. BIGGINS

University of Sheffield

In a discrete-time supercritical branching random walk, let $Z^{(n)}$ be the point process formed by the nth generation. Let $m(\lambda)$ be the Laplace transform of the intensity measure of $Z^{(1)}$. Then $W^{(n)}(\lambda) = \int e^{-\lambda x} Z^{(n)}(dx)/m(\lambda)^n$, which is the Laplace transform of $Z^{(n)}$ normalized by its expected value, forms a martingale for any λ with $|m(\lambda)|$ finite but nonzero. The convergence of these martingales uniformly in λ, for λ lying in a suitable set, is the first main result of this paper. This will imply that, on that set, the martingale limit $W(\lambda)$ is actually an analytic function of λ. The uniform convergence results are used to obtain extensions of known results on the growth of $Z^{(n)}(nc + D)$ with n, for bounded intervals D and fixed c. This forms the second part of the paper, where local large deviation results for $Z^{(n)}$ which are uniform in c are considered. Finally, similar results, both on martingale convergence and uniform local large deviations, are also obtained for continuous-time models including branching Brownian motion.

1. Introduction. An initial ancestor is at the origin (in $\mathscr{R}^p$) and the positions of its children form a point process (the first generation) $Z^{(1)}$, with intensity measure μ. Each of these has children, the second generation, with positions relative to their parent given by independent copies of $Z^{(1)}$, and so on. The point process formed by the nth generation is denoted by $Z^{(n)}$, with points $\{z_r^{(n)}: r\}$, and this has the intensity measure μ^{n*}, the n-fold convolution of μ. The (multivariate) Laplace transform of μ is denoted by m, so that

$$m(\lambda) = \int e^{-\lambda x} \mu(dx),$$

where $\lambda \in \mathscr{C}^p$ and $\lambda = \theta + i\eta$ with $\theta, \eta \in \mathscr{R}^p$. (Throughout, the real and imaginary parts of λ will be denoted by θ and η, respectively; also, no special notation will be used to indicate inner products like λx above.) Sometimes we need to consider the transform m as a function of θ, with $\eta = 0$, and this will be denoted by $m(\theta)$; while $m(i\eta)$ is $m(\lambda)$ with $\lambda = i\eta$. We will throughout consider only those λ for which $m(\theta) < \infty$. Notice that $m(0) = \mu(\mathscr{R}^p) = E(Z^{(1)}(\mathscr{R}^p))$ is a parent's expected number of children; we will consider only supercritical processes and so take this to be greater than 1.

Let $\mathscr{F}^{(n)}$ be the σ-field containing all information about the first n generations. It is straightforward to show, and well known, that, when $m(\lambda)$

Received January 1990; revised October 1990.
AMS 1980 *subject classifications*. Primary 60J80; secondary 60F10, 60G42, 60G44.
Key words and phrases. Spatial growth in branching processes, uniform local large deviations, Banach space valued martingales.

exists and is finite but nonzero,

$$W^{(n)}(\lambda) = m(\lambda)^{-n} \int e^{-\lambda x} Z^{(n)}(dx)$$

is a martingale with respect to $\mathscr{F}^{(n)}$. The convergence of these or similar martingales has been considered quite frequently [see Watanabe (1967), Joffe, Le Cam and Neveu (1973), Kingman (1975), Biggins (1977), Wang (1980), Uchiyama (1982) and Neveu (1988)]. When this martingale converges we will denote its limit by $W(\lambda)$. The first objective here is to seek conditions ensuring $W^{(n)}(\lambda)$ converges to $W(\lambda)$ uniformly for λ in any closed subset of a certain open set, Λ, almost surely. As $W^{(n)}(\lambda)$ is analytic on Λ, it will follow that $W(\lambda)$ will also be analytic on the interior of Λ.

In another paper [Biggins (1989)], I considered the problem in the one-dimensional case (i.e., $p = 1$) and obtained convergence in a neighbourhood of part of the real axis. Here greater efforts are made to make the set Λ as large as possible with an explicit definition. The idea used there is to employ Cauchy's integral formula to estimate the supremum of $|W^{(n+r)}(\lambda) - W^{(n)}(\lambda)|$ over a set strictly within a contour in terms of an integral round the contour and then to show that this integral tends to 0 with n, uniformly in r. This method extends to the multidimensional framework. Also in that paper is a result on the uniform convergence of $W^{(n)}(\theta)$ for real θ, which can be obtained under rather weaker moment conditions. Of course, the trick of using Cauchy's integral formula is then no longer available. It ought to be possible to extend this result also to the multidimensional framework, but this problem is not considered here. The only explicit result I know of on such uniform convergence is contained in Joffe, Le Cam and Neveu (1973), who consider the convergence of $W^{(n)}(i\eta)$ for a particular case of the process described above.

A major reason for trying to obtain a uniform convergence result for $W^{(n)}(\lambda)$ (besides its intrinsic interest) is to use it in obtaining results about the sequence of point processes $\{Z^{(n)}\}$ as $n \to \infty$. It is plausible that $Z^{(n)}$ will have many properties analogous to those of its intensity measure, the n-fold convolution μ^{n*}, and as transform methods are frequently used to establish properties of the latter, it is reasonable to hope that sufficiently strong results about the martingale $W^{(n)}(\lambda)$ will be useful in establishing results for $Z^{(n)}$. This general idea is not new [see, e.g., Watanabe (1967) and Remark 2 in Joffe and Moncayo (1973)] but, as I hope to demonstrate both here and elsewhere, it still has considerable potential.

Here we will use $W^{(n)}(\lambda)$ in the study of the large deviation behaviour of $Z^{(n)}$. (Note that here "large deviation behaviour for $Z^{(n)}$" refers to the extent to which large deviation estimates for μ^{n*} carry over to $Z^{(n)}$, rather than to the asymptotics of probabilities of rare events associated with $Z^{(n)}$.) Stone (1967) gives an estimate of $\mu^{n*}(nc + dx)$ that is uniform in c, for a suitable range of values of c—a uniform local large deviation theorem. The main result in the later part of the paper will be (roughly speaking) a similar uniform estimate of $Z^{(n)}(nc + dx)$. This estimate will involve $W(\theta)$ in an essential way. Some weaker results on the large deviations of $Z^{(n)}$ have been established previously. In Biggins (1979), Corollary 1 and Theorem B, it is shown that (when $p = 1$)

$$\frac{Z^{(n)}(nc + a, nc + b)}{EZ^{(n)}(nc + a, nc + b)} \to W(\theta) \tag{1.1}$$

almost surely, for any (fixed) c in a suitable interval, where θ is given by $-m'(\theta)/m(\theta) = c$. [If μ is lattice (a, b) must be large enough to contain a lattice point.] A result of the same form for a multivariate continuous-time model is given by Uchiyama (1982), Theorem 1 and Remark 3. It will be a consequence of Corollary 4 in Section 4 that (1.1) holds uniformly, for c in suitable compact subsets, almost surely. It is not surprising that the uniform convergence of $W^{(n)}$ to W and the smoothness of the limit W should be essential ingredients in establishing this.

The next section contains the statement of the main results on the convergence of $W^{(n)}(\lambda)$ for this discrete-time model, while Section 3 contains their proofs. Section 4 contains both the statement and the proofs of the large deviation results for $Z^{(n)}$. The final section discusses the continuous-time analogues of the martingale convergence and large deviation results, essentially demonstrating that they carry over with only the obvious changes. The model discussed in that section includes as special cases both branching Brownian motion and that studied by Uchiyama (1982). In particular, the large deviation results give a generalization of Uchiyama's Theorem 1, mentioned above.

2. Martingale convergence, discrete time; results.

The first theorem considers the convergence of $W^{(n)}(\lambda)$ for a particular λ, rather than as a function of λ. The convention that $\lambda = \theta + i\eta$, with $\theta, \eta \in \mathscr{R}^p$, established earlier, is used in its statement. Its two corollaries deal with the special cases $\lambda = \theta$ and $\lambda = i\eta$. The set Ω^0 occurring in the first of these, is defined by $\Omega^0 = \text{int}\{\lambda\colon m(\theta) < \infty\}$, where int A is the interior of the set A. Also $m'(\theta)$ is a vector of partial derivatives with respect to θ, so that $\theta m'(\theta)$ is an inner product.

THEOREM 1. *If*

$$EW^{(1)}(\theta)^{\gamma} < \infty \quad \textit{for some } \gamma \in (1, 2] \tag{2.1}$$

and

$$\frac{m(\alpha\theta)}{|m(\lambda)|^{\alpha}} < 1 \quad \textit{for some } \alpha \in (1, \gamma], \tag{2.2}$$

then $\{W^{(n)}(\lambda)\}$ *converges almost surely and in* α*th mean.*

COROLLARY 1. *If* (2.1) *holds*, $\theta \in \Omega^0$ *and*

$$-\log(m(\theta)) < -\frac{\theta m'(\theta)}{m(\theta)}, \tag{2.3}$$

then $\{W^{(n)}(\theta)\}$ *converges in mean.*

COROLLARY 2. *If* (2.1) *holds and*

$$m(0) < |m(i\eta)|^{\gamma},$$

then $\{W^{(n)}(i\eta)\}$ *converges almost surely and in* γ*th mean.*

A conclusion like that in Theorem 1 has been obtained for a continuous-time version of the process by Uchiyama (1982), Proposition 1. In Section 5, Theorem 1 will be used to obtain a strengthening of that result.

In the case $\lambda = \theta$, $W^{(n)}(\theta)$ is a positive martingale so its convergence almost surely is automatic. This case has been considered in greater detail in Biggins (1977) [see also Biggins (1978), page 71], where, when $\theta \in \Omega^0$, the weaker moment condition $EW^{(1)}(\theta)\log^+(W^{(1)}(\theta)) < \infty$ and (2.3) are shown to be necessary and sufficient for $W^{(n)}(\theta)$ to converge in mean. Also, for real θ, the conditions of Theorem 1 are essentially necessary for convergence in αth mean; see Biggins (1979), page 26.

Turning to the case $\lambda = i\eta$ covered in Corollary 2, here $m(0)$ is the mean family size and (2.1) is just a moment condition on the family size. Some results of this kind have been obtained previously, at least when $\gamma = 2$. See, for example, Stam (1966), in particular equation (16), and Joffe, Le Cam and Neveu (1973), who also consider uniform convergence.

We now turn to the uniform convergence of $\{W^{(n)}(\lambda)\}$. Let

$$\Omega_\gamma^1 = \operatorname{int}\{\lambda \colon EW^{(1)}(\theta)^{\gamma} < \infty\}$$

and

$$\Omega_\gamma^2 = \operatorname{int}\left\{\lambda \colon \frac{m(\gamma\theta)}{|m(\lambda)|^{\gamma}} < 1\right\}. \tag{2.4}$$

Now let

$$\Omega_\gamma = \Omega_\gamma^1 \cap \Omega_\gamma^2 \quad \text{and} \quad \Lambda = \bigcup_{1<\gamma\le 2} \Omega_\gamma .$$

Therefore Ω_γ and Λ are open.

THEOREM 2. $\{W^{(n)}(\lambda)\}$ *converges uniformly on any compact subset of* Λ, *almost surely and in mean, as* $n \to \infty$.

COROLLARY 3. $W(\lambda)$ *is analytic on* Λ.

If $W^{(n)}$ is restricted to a compact subset $F \subset \Lambda$, it can be thought of as a martingale with values in the Banach space of continuous functions on F

(under the supremum norm). Theorem 2 then asserts that this martingale converges almost surely and in mean. One possible approach to this result is to show that $W(\lambda)$, the limit of $W^{(n)}(\lambda)$, which exists by Theorem 1, is continuous in λ, and then to use known results about the convergence of $W^{(n)} = E(W|\mathscr{F}^{(n)})$ to W in Banach spaces [Neveu (1975), Proposition V-2-6]. This is the approach adopted by Joffe, Le Cam and Neveu (1973) in obtaining their result of this kind. Here a direct approach is used and so the continuity of W on F is a consequence of the theorem. In fact, as $W^{(n)}$ is analytic on Λ, standard complex analysis [see Hörmander (1973), Corollary 2.2.4] gives the analyticity of W recorded in Corollary 3.

Some remarks about the constituent sets of Λ are in order. Obviously, Ω_γ^1 is a "strip," in that $\theta + i\eta$ is in Ω_γ^1 if and only if θ is, and the same is true of Ω^0. As

$$(m(\theta)W^{(1)}(\theta))^\gamma = \left(\sum e^{-\theta z_r^{(1)}}\right)^\gamma \tag{2.5}$$

is convex in θ for $\gamma \geq 1$, so is $m(\theta)^\gamma EW^{(1)}(\theta)^\gamma$; hence $\{\theta\colon \theta \in \Omega^0\}$ and $\{\theta\colon \theta \in \Omega_\gamma^1\}$ are both open convex sets in $\mathscr{R}^p$. It is also worth noting here that $m(\lambda)$, $EW^{(1)}(\theta)^\gamma$ and $m(\gamma\theta)/|m(\lambda)|^\gamma$ are continuous functions on Ω^0, Ω_γ^1 and Ω_γ^2, respectively.

To clarify the role of γ it is worth re-expressing slightly the definition of Λ. Let

$$\Omega_\gamma^3 = \left\{\lambda\colon \lambda \in \Omega^0,\ \inf_{1 \leq \alpha \leq \gamma} \frac{m(\alpha\theta)}{|m(\lambda)|^\alpha} < 1\right\}.$$

Then it is fairly easy to check that Ω_γ^3 can replace Ω_γ^2 in the definition of Λ so that Λ is also given by

$$\Lambda = \bigcup_{1 < \gamma \leq 2} \left(\Omega_\gamma^1 \cap \Omega_\gamma^3\right).$$

Now notice that Ω_γ^1 increases as γ decreases but Ω_γ^3 decreases as γ decreases, so that there is a trade-off between the moment condition, Ω_γ^1, and the condition on $m(\lambda)$, Ω_γ^3. Also the restriction of Ω_γ^3 to $\theta \in \mathscr{R}^p$ is

$$\left\{\theta\colon \theta \in \Omega^0,\ \inf_{1 \leq \alpha \leq \gamma} \frac{m(\alpha\theta)}{m(\theta)^\alpha} < 1\right\}$$

and, by considering the slope of $m(\alpha\theta)/m(\theta)^\alpha$ at $\alpha = 1$, this is the same as

$$\left\{\theta\colon \theta \in \Omega^0,\ -\log(m(\theta)) < -\frac{\theta m'(\theta)}{m(\theta)}\right\},$$

and so is actually independent of γ. (The same calculation is all that is needed to deduce Corollary 1 from Theorem 1.)

3. Martingale convergence, discrete time; proofs. The basic ideas here are the same as those used to prove Theorem 1 in Biggins (1989). The definition of the process gives immediately that

$$(3.1) \qquad W^{(n+1)}(\lambda) - W^{(n)}(\lambda) = \sum_r \frac{e^{-\lambda z_r^{(n)}}}{m(\lambda)^n}(W_{1,r}(\lambda) - 1),$$

where $\{W_{1,r}(\lambda)\}$ are i.i.d. copies of $W^{(1)}(\lambda)$. Hence the martingale difference $W^{(n+1)}(\lambda) - W^{(n)}(\lambda)$ is, given $\mathscr{F}^{(n)}$, the weighted sum of independent identically distributed random variables with zero mean. The following lemma concerns the calculation of the moments of such expressions. A stronger result for real random variables is contained in von Bahr and Esseen (1965), Theorem 2, and their methods yield the following result for complex-valued random variables fairly easily. The actual value of the constant 2^α occurring is irrelevant, so the Marcinkiewicz–Zygmund inequalities and their extension to martingales [see, e.g., Neveu (1975), Corollary VIII-3-18] will also readily yield a suitable inequality.

LEMMA 1. *If $\{X_i\}$ are independent complex random variables with $E(X_i) = 0$, or more generally, martingale differences, then $E|\sum X_i|^\alpha \le 2^\alpha \sum E|X_i|^\alpha$ for $1 \le \alpha \le 2$.*

PROOF. Let $S_n = \sum^n X_i$ and let X'_{n+1} be independent of $\{X_1, X_2, \dots, X_n\}$ with the same distribution as X_{n+1}. Then $E|S_{n+1}|^\alpha = E|S_n + X_{n+1}|^\alpha \le E|S_n + X_{n+1} - X'_{n+1}|^\alpha \le E|S_n|^\alpha + E|X_{n+1} - X'_{n+1}|^\alpha \le E|S_n|^\alpha + 2^\alpha E|X_{n+1}|^\alpha$, where the inequalities used are, successively, those of Jensen, Clarkson [see von Bahr and Esseen (1965)] and Minkowski. □

An easy calculation establishes that

$$(3.2) \qquad E|W^{(1)}(\lambda) - 1|^\alpha \le 2^{\alpha+1}\frac{m(\theta)^\alpha}{|m(\lambda)|^\alpha} EW^{(1)}(\theta)^\alpha$$

and this estimate and Lemma 1 now allow the following estimates of moments of $(W^{(n+1)} - W^{(n)})$ and hence of $(W^{(n)} - 1)$. Notice that Theorem 1 is an immediate consequence of this lemma.

LEMMA 2. *For $1 < \alpha \le 2$ and fixed λ, let $\kappa = m(\alpha\theta)/|m(\lambda)|^\alpha$, $M = EW^{(1)}(\theta)^\alpha$ and $\phi = m(\theta)/|m(\lambda)|$ then*

$$\text{(i)} \qquad E|W^{(n+1)} - W^{(n)}|^\alpha \le 2^{2\alpha+1}\phi^\alpha M \kappa^n,$$

$$\text{(ii)} \qquad E|W^{(n+1)} - 1|^\alpha \le 2^{3\alpha+1}\phi^\alpha M \sum_{r=0}^{n} \kappa^r,$$

$$\text{(iii)} \qquad \sum E|W^{(n+1)} - W^{(n)}| \le \frac{2^3\phi M^{1/\alpha}}{1 - \kappa^{1/\alpha}} \quad \text{if } \kappa < 1.$$

PROOF. Take the expectation of the αth absolute moment of (3.1) conditional on $\mathscr{F}^{(n)}$, apply Lemma 1, the bound (3.2) and take unconditional expectations; this gives (i). The other two parts follow easily. □

PROOF OF THEOREM 2. The proof of this theorem will rely on an estimate derived using Cauchy's integral formula for functions of several complex variables [see, e.g., Hörmander (1973), Theorem 2.2.1]. The theory required (and the derivation of the estimate) is a straightforward generalisation of the one-variable case. To state the estimate, we first introduce some notation.

The open polydisc centred at $x = (x_1, x_2, \dots, x_p) \in \mathscr{C}^p$ with radius $\rho > 0$ is denoted by $D_x(\rho)$ and defined by $D_x(\rho) = \{y \in \mathscr{C}^p: |x_j - y_j| < \rho, \forall j\}$, and its "distinguished boundary" $\Gamma_x(\rho)$ is defined by $\Gamma_x(\rho) = \{y \in \mathscr{C}^p: |x_j - y_j| = \rho, \forall j\}$. Now suppose $\Gamma_x(2\rho)$ is parameterized by $t \in C \subset \mathscr{R}^p$, where

$$(3.3) \qquad C = \{t \in \mathscr{R}^p: 0 \le t_j \le 2\pi, \forall j\} \quad \text{and} \quad z_j(t) = x_j + 2\rho e^{it_j},$$

so that $\Gamma_x(2\rho) = \{z(t): t \in C\}$.

LEMMA 3. *If f is analytic on $D_x(2\rho')$ with $\rho' > \rho$, then*

$$\sup_{\lambda \in D_x(\rho)} |f(\lambda)| \le \pi^{-p} \int_C |f(z(t))|\, dt$$

with C and $z(t)$ as defined at (3.3).

PROOF. Use Cauchy's integral formula over $\Gamma_x(2\rho)$ and the triangle inequality. □

LEMMA 4. *For any $x \in \Lambda$ there is a polydisc $D_x(\rho) \subset \Lambda$ such that $W^{(n)}(\lambda)$ converges uniformly on $D_x(\rho)$, almost surely and in mean.*

PROOF. Given any $x \in \Lambda$, we can find γ with $x \in \Omega_\gamma$, and hence can find ρ such that $D_x(3\rho) \subset \Omega_\gamma$. As $W^{(N+1)}(\lambda) - W^{(n)}(\lambda)$ is analytic in λ on $D_x(3\rho)$, we may use the estimate in Lemma 3 to deduce that

$$(3.4) \qquad \sup_{N \ge n} \sup_{\lambda \in D_x(\rho)} \pi^p |W^{(N+1)}(\lambda) - W^{(n)}(\lambda)| \le \int_C \sum_{r=n}^{\infty} |W^{(r+1)} - W^{(r)}|\, dt$$

[where $z(t)$ has been suppressed in the integrand]. Note that [with $\Gamma = \Gamma_x(2\rho)$]

$$E\int_C \sum_{r=0}^{\infty} |W^{(r+1)} - W^{(r)}|\, dt \le (2\pi)^p \sup_{\lambda \in \Gamma} \sum_{r=0}^{\infty} E|W^{(r+1)}(\lambda) - W^{(r)}(\lambda)|,$$

so if this bound is finite the right-hand side of (3.4) will go to 0 both in mean and almost surely. Recall that $\Gamma \subset \Omega_\gamma = \Omega_\gamma^1 \cap \Omega_\gamma^2$, Γ is closed and compact and $m(\lambda)$, $EW^{(1)}(\theta)^\gamma$ and $m(\gamma\theta)/|m(\lambda)|^\gamma$ are continuous on Ω_γ and hence on Γ. Therefore the estimate in Lemma 2(iii), with $\alpha = \gamma$, is uniformly bounded for $\lambda \in \Gamma$, completing the proof of Lemma 4 and hence, by a compactness argument, of Theorem 2. □

4. Large deviations for $Z^{(n)}$. In this section we will obtain large deviation results for $Z^{(n)}$ showing that, up to the random function $W(\theta)$, it mimics the behaviour of μ^{n*}. The relevant estimates for μ^{n*} are supplied by Stone (1967) and the results for $Z^{(n)}$ will be obtained by following Stone's proof through with $Z^{(n)}$ in place of μ^{n*}. We will assume that μ is nonlattice, in that for some (and then for all) $\theta \in \Omega^0$, $|m(\theta + i\eta)/m(\theta)| = 1$ only when $\eta = 0$. This assumption is unnecessary, as a study of Stone's proof will reveal, but it lessens the notational burden. We will also assume that μ is strictly p-dimensional in that its support is not contained in any lower-dimensional hyperplane.

Let the conjugate measure μ_θ be given by

$$\mu_\theta(dx) = \frac{e^{-\theta x}}{m(\theta)}\mu(dx),$$

with mean c_θ and covariance Σ_θ. Let σ_θ be the square root of the determinant of Σ_θ, which is nonzero as μ_θ is strictly p-dimensional, and let p_θ be the Gaussian density with covariance Σ_θ. Let I_h be the cube of side h centred at the origin. Finally, if $A \subset \mathscr{C}^p$ let $\tilde{A}$ be the restriction of A to $\theta \in \mathscr{R}^p$, so, for example, $\tilde{\Omega}^0 = \operatorname{int}\{\theta \in \mathscr{R}^p\colon\ m(\theta) < \infty\}$. The following result is (essentially) contained in Theorem 2 of Stone (1967).

THEOREM 3.

$$\left| n^{p/2}\mu_\theta^{n*}(x + nc_\theta + I_h) - h^p p_\theta\left(\frac{x}{\sqrt{n}}\right)\right| \to 0$$

as $n \to \infty$, uniformly for h in bounded sets, θ in compact subsets of $\tilde{\Omega}^0$ and all x.

There are two parts to Stone's proof. First, the result is proved for a smoothed version of μ_θ^{n*}. This is done (essentially) by estimating the required difference by the integral of the absolute difference in the corresponding characteristic functions. The second part then shows that the smoothed and unsmoothed versions of μ_θ^{n*} attach similar weight to sets of the form $x + I_h$. (When there is a lattice component it is not smoothed, the Fourier estimate alone sufficing for these parts.)

To state the analogue of this result for the branching random walk, let

$$Z_\theta^{(n)}(dx) = \frac{e^{-\theta x}}{m(\theta)^n}Z^{(n)}(dx).$$

THEOREM 4.

$$\left| n^{p/2}Z_\theta^{(n)}(x + nc_\theta + I_h) - h^p W(\theta) p_\theta\left(\frac{x}{\sqrt{n}}\right)\right| \to 0$$

as $n \to \infty$, uniformly for h in bounded sets, θ in compact subsets of $\tilde{\Lambda}$ and all x, almost surely.

We will discuss the proof in a moment but note first that

$$\int f(y)Z^{(n)}(nc_\theta + dy) = \left(m(\theta)e^{\theta c_\theta}\right)^n \int f(y)e^{\theta y} Z_\theta^{(n)}(nc_\theta + dy).$$

We can therefore use a Riemann sum argument and Theorem 4 to approximate this integral. By paying suitable attention to the uniformity in θ of the Riemann approximation of $f(y)e^{\theta y}$, this yields uniform estimates for $Z^{(n)}$ [cf. Theorem 3 of Stone (1967)]. We give one version; others are possible.

COROLLARY 4. *For any directly Riemann integrable function f of compact support,*

$$\frac{n^{p/2}}{\left(m(\theta)e^{\theta c_\theta}\right)^n} \int f(y)Z^{(n)}(nc_\theta + dy) \to \frac{W(\theta)}{(2\pi)^{p/2}\sigma_\theta} \int f(y)e^{\theta y}\, dy$$

as $n \to \infty$, *uniformly for* θ *in compact subsets of* $\tilde{\Lambda}$, *almost surely.*

Notice that if f is the indicator function of a bounded Riemann integrable set D, then this theorem gives an estimate of the growth of $Z^{(n)}(nc_\theta + D)$ that is uniform in θ. Combining this with the corresponding estimate of $\mu^{n*}(nc_\theta + D)$, derived from Theorem 3, gives a version of (1.1) holding uniformly in θ.

The function $-m'(\theta)/m(\theta)$ is continuous and one to one on $\tilde{\Omega}^0$, so we see that compact subsets of $\tilde{\Lambda}$ correspond to compact subsets of $C_{\tilde{\Lambda}} = \{c: \theta \in \tilde{\Lambda}, c = -m'(\theta)/m(\theta)\}$. The expressions in Theorem 4 and Corollary 4 could just as well be parameterized in terms of c (instead of θ) with the corresponding results holding uniformly in compact subsets of $C_{\tilde{\Lambda}}$. It is worth commenting that, under this parameterization, the term $m(\theta)e^{\theta c_\theta}$ in the normalization in Corollary 4 is exactly the Cramér function of μ, defined by $\inf\{m(\theta)e^{\theta c}: \theta\}$, evaluated at c.

PROOF OF THEOREM 4. Let $\phi_\theta(\eta) = \int e^{-i\eta x}\mu_\theta(dx)$ be the characteristic function of μ_θ (the choice of $-\eta$ in the exponent is for later notational convenience). We will also sometimes denote this by $\phi(\lambda)$ $[= m(\lambda)/m(\theta)]$. The "characteristic function" of $Z_\theta^{(n)}$ is then $W^{(n)}(\theta + i\eta)\phi_\theta(\eta)^n$ $[= W^{(n)}(\lambda)\phi(\lambda)^n$, and should be interpreted as $m(\theta)^{-n}\int e^{-\lambda x}Z^{(n)}(dx)$ if $m(\lambda) = 0]$. The next lemma shows that the approximation of $W^{(n)}(\lambda)\phi(\lambda)^n$ by $W(\theta)\phi(\lambda)^n$ is legitimate. Once the lemma is established, the theorem follows from the proof of Theorem 3.

LEMMA 5. *Let F be a compact subset of* $\tilde{\Lambda}$. *Then, for any* $a > 0$,

$$\sup_{\theta \in F} \int_{|\eta| \le a} n^{p/2} \left| W^{(n)}(\theta + i\eta) - W(\theta) \right| \left| \phi_\theta(\eta)^n \right| d\eta \to 0$$

as $n \to \infty$ *almost surely, with the null set independent of a.*

PROOF. We split the integral into two parts, $|\eta| < \varepsilon$ and $\varepsilon \le |\eta| \le a$, and deal with them in that order.

A standard Taylor series estimation gives, for small ε and $|\eta| \le \varepsilon$,

$$\sup_{\theta \in F} |\phi_\theta(\eta)| \le \exp(-C|\eta|^2).$$

Now let $F_\varepsilon = \{\lambda\colon \theta \in F, |\eta| \le \varepsilon\}$. The integral over $|\eta| < \varepsilon$ is less than

$$\sup_{\lambda \in F_\varepsilon} |W^{(n)}(\lambda) - W(\theta)| \int_{|\eta| < \varepsilon\sqrt{n}} \exp(-C|\eta|^2)\, d\eta \to K \sup_{\lambda \in F_\varepsilon} |W(\lambda) - W(\theta)|,$$

using Theorem 2, and this may be made arbitrarily small by choosing ε sufficiently small.

Consider now

$$\int_{\varepsilon \le |\eta| \le a} n^{p/2} |W^{(n)}(\lambda)\phi(\lambda)^n|\, d\eta.$$

We will show that this converges to 0 uniformly in a neighbourhood of any $\theta_0 \in \tilde{\Lambda}$, and hence uniformly on F, completing the proof.

Let B_r be a closed ball in $\mathscr{R}^p$ of radius r centred at $\theta_0 \in \tilde{\Lambda}$ and let $G_{br} = \{\lambda\colon \theta \in B_{br}, b^{-1}\varepsilon \le |\eta| \le ba\}$ so that $G_r \subset G_{2r}$. As $\theta_0 \in \Lambda$, we may take $\gamma \in (1, 2]$ such that $\theta_0 \in \Omega_\gamma$ and then

$$\frac{m(\gamma\theta_0)^{1/\gamma}}{m(\theta_0)} < 1. \tag{4.1}$$

We may therefore choose r sufficiently small that, for some $\delta < 1$, $B_{2r} \subset \Omega_\gamma$,

$$\frac{\sup\{m(\gamma\theta)^{1/\gamma}\colon \theta \in B_{2r}\}}{\inf\{m(\theta)\colon \theta \in B_{2r}\}} \le \delta \tag{4.2}$$

and, as μ is nonlattice,

$$\frac{\sup\{|m(\lambda)|\colon \lambda \in G_{2r}\}}{\inf\{m(\theta)\colon \theta \in B_{2r}\}} \le \delta. \tag{4.3}$$

Let $B^{(n)}(\lambda) = \int e^{-\lambda x} Z^{(n)}(dx) - m(\lambda)^n$ so that

$$W^{(n)}(\lambda)\phi(\lambda)^n = m(\theta)^{-n} B^{(n)}(\lambda) + \phi(\lambda)^n,$$

then

$$\begin{aligned} &\sup_{\theta \in B_r} n^{p/2} \int_{\varepsilon \le |\eta| \le a} |W^{(n)}(\lambda)\phi(\lambda)^n|\, d\eta \\ &\qquad \le K n^{p/2} \left\{ \frac{\sup\{|B^{(n)}(\lambda)|\colon \lambda \in G_r\}}{\inf\{m(\theta)^n\colon \theta \in B_r\}} + \delta^n \right\}, \end{aligned} \tag{4.4}$$

where here, and in what follows, K is a suitable constant independent of n and θ. As $B^{(n)}(\lambda)$ is analytic in λ, we can use Lemma 3 and a compactness argument to show that

$$(4.5)\qquad \sup_{\lambda \in G_r} |B^{(n)}(\lambda)| \le \pi^{-p} \sum_j \int_{C_j} |B^{(n)}(z(s))|\, ds,$$

where $\{C_j\}$ parameterize the distinguished boundaries of a finite number of polydiscs comfortably covering G_r and lying within G_{2r}. ("Comfortably" here means that if their radii were halved they would still cover G_r.) Taking expected values of (4.5) and using Jensen's inequality, we see that, for $\alpha > 1$,

$$(4.6)\qquad E \sup_{\lambda \in G_r} |B^{(n)}(\lambda)| \le K \sup_{\lambda \in G_{2r}} \left(E|B^{(n)}(\lambda)|^{\alpha}\right)^{1/\alpha}.$$

The next lemma, which is essentially a restatement of Lemma 2(ii), bounds the right-hand side here.

LEMMA 6. *For* $1 < \alpha \le 2$,

$$E|B^{(n+1)}(\lambda)|^{\alpha} \le 2^{3\alpha+1} m(\theta)^{\alpha} EW^{(1)}(\theta)^{\alpha} \sum_{r=0}^{n} m(\alpha\theta)^{r} |m(\lambda)|^{(n-r)\alpha}.$$

As $B_{2r} \subset \Omega^1_{\gamma}$, $\sup\{m(\theta)^{\gamma} EW^{(1)}(\theta)^{\gamma}\colon \theta \in B_{2r}\}$ is finite. Now take $\alpha = \gamma$ in (4.6) and use it, Lemma 6, (4.2) and (4.3) to see that the expected value of the right-hand side of (4.4) is less than

$$Kn^{p/2}\left(\left(\sum_{r=0}^{n-1} \delta^{\alpha r}\delta^{\alpha(n-1-r)}\right)^{1/\alpha} + \delta^{n}\right) \le Kn^{p/2}(n^{1/\alpha} + 1)\delta^{n-1}.$$

Consequently, the right-hand side of (4.4) converges to 0 almost surely as $n \to \infty$, and, as the expression is increasing in a, the null set can be taken independent of a. This completes the proof of the lemma. □

In following the pattern of Stone's proof of Theorem 3, we find we must estimate [cf. Stone's (3.6) and (3.8)]

$$\int_{|\eta| \le a} n^{p/2} e^{ix\eta} W^{(n)}(\theta + i\eta) \phi_\theta(\eta)^{n} g(\eta)\, d\eta,$$

where g depends on a and h but satisfies $|g(\eta)| \le 1$. Obviously, Lemma 5 allows us to instead estimate

$$\int_{|\eta| \le a} n^{p/2} e^{ix\eta} W(\theta) \phi_\theta(\eta)^{n} g(\eta)\, d\eta,$$

and this differs from the quantity considered by Stone only by the factor $W(\theta)$ (which is continuous in θ). □

5. Continuous-time results. Suppose that each particle lives for an exponentially distributed length of time (hazard rate β) and during its lifetime it moves according to an independent copy of a process with stationary independent increments. This process is taken to be regular, in the sense that it has right-continuous sample paths with limits from the left, and its exponent function is $a(\lambda)$, so that if a particle, initially at the origin, is still alive at time t, then its position has transform $\exp(a(\lambda)t)$. At death a particle is replaced by its children, and the point process X gives their positions relative to their parent's at its death. Independent copies of X are, of course, used for each family, and we let $b(\lambda) = E\int e^{-\lambda x}X(dx)$. Let $Z^{(t)}$ be the point process resulting at time t from this construction. The total number of particles $\{Z^{(t)}(\mathscr{R}^p)\}$ forms a Markov branching process, and we assume that this is nonexplosive. If $a(\lambda)$ is degenerate, so that a particle does not move during its lifetime, this is the process discussed by Uchiyama (1982). If movement is according to a homogeneous Brownian motion and X is concentrated at 0, we have branching Brownian motion, while if instead the branching process is a birth-and-death process and the displacement distribution for an offspring is isotropic, then we have a process discussed by Wang (1980).

The results in Section 2 may be applied to any discrete skeleton of this process, thereby allowing continuous-time versions of these theorems to be established. However, this involves imposing conditions on $Z^{(t)}$ for all small t, and it is more usual to express such conditions in terms of the "infinitesimal" conditions. One part of this is fairly straightforward, for, by conditioning on the time of the first death, it follows easily that

$$m(\lambda)^t = E\int e^{-\lambda x}Z^{(t)}(dx) = \exp((\beta b(\lambda) - \beta + a(\lambda))t), \tag{5.1}$$

so Ω_γ^2, defined by (2.4), is independent of t, and could easily be given a formulation in terms of a, b and β.

The moment condition presents more problems, but it is possible by a fairly straightforward but tedious calculation to show that $E(\int e^{-\theta x}Z^{(t)}(dx))^\gamma$ is finite for all small t, provided that $a(\theta\gamma)$, $b(\theta\gamma)$ and $E(\int e^{-\theta x}X(dx))^\gamma$ are all finite. Consequently, for the analogue of Theorem 2 we can now take

$$\Omega_\gamma^1 = \operatorname{int}\left\{\lambda\colon a(\theta\gamma) < \infty,\ b(\theta\gamma) < \infty,\ E\left(\int e^{-\theta x}X(dx)\right)^\gamma < \infty\right\}$$

and let Ω_γ and Λ be defined as before. Note that Λ is again automatically open.

The process $\{Z^{(t)}\}$ can be taken to be regular, with respect to the topology of weak convergence of measures. Then $\int e^{-\lambda x}Z^{(t)}(dx)$ will be regular so if Theorem 1 holds for any discrete skeleton, it also holds for $W^{(t)}(\lambda)$. Hence we have the following result.

THEOREM 5. *If $a(\theta\gamma)$, $b(\theta\gamma)$ and $E(\int e^{-\theta x}X(dx))^\gamma$ are all finite for some $\gamma \in (1,2]$ and*

$$\frac{m(\alpha\theta)}{|m(\lambda)|^\alpha} < 1 \tag{5.2}$$

for some $\alpha \in (1,\gamma]$, then $\{W^{(t)}(\lambda)\}$ converges almost surely and in αth mean.

A result of this kind is given as Proposition 1 of Uchiyama (1982) and a special case is covered by Proposition 1 of Neveu (1988). To discuss the relationship of Uchiyama's result with Theorem 5, let $T = \{\theta\colon m(\alpha\theta)/m(\theta)^{\alpha} < 1 \text{ for some } \alpha > 1\}$. This is simply a reformulation of the definition of T given by Uchiyama (1982), just before (2.4). Now let $\hat{\Omega}_{\alpha} = \{\lambda\colon \alpha\theta \in T, m(\alpha\theta)/|m(\lambda)|^{\alpha} < 1\}$, then the condition (2.4) imposed by Uchiyama is essentially: $\lambda \in \hat{\Omega}_{\alpha}$ for some $\alpha \in (1,2]$. In addition, the moment condition, labelled (A.1) there,

$$E\int X(dy)\int e^{-\theta x}X(dx) < \infty \quad \text{for all } \theta \in T, \tag{5.3}$$

is imposed. Proposition 1 of Uchiyama states that these two conditions suffice for the conclusion of Theorem 5. The next result shows that this result is contained in Theorem 5. In fact, when $0 \in T$, (5.3) includes the condition that the second moment of the family size is finite, so Uchiyama's conditions are stronger than those employed here.

Proposition 1. *If* (5.3) *holds and* $\lambda \in \hat{\Omega}_{\gamma}$ *with* $1 < \gamma \leq 2$, *then the conditions of Theorem* 5 *hold with* $\alpha = \gamma$.

Proof. If $\lambda \in \hat{\Omega}_{\gamma}$, then $\gamma\theta \in T$ so, from (5.3), $E\int X(dy)\int e^{-\gamma\theta x}X(dx) < \infty$ and, letting $N = \int X(dy)$, $E(\int e^{-\theta x}X(dx))^{\gamma} \leq EN^{\gamma-1}\int e^{-\gamma\theta x}X(dx) \leq EN\int e^{-\gamma\theta x}X(dx) < \infty$. Furthermore, again as $\gamma\theta \in T$, $m(\gamma\theta) < \infty$ so $b(\gamma\theta)$ and $a(\gamma\theta)$ are finite. Of course, $\lambda \in \hat{\Omega}_{\gamma}$ also implies that (5.2) holds with $\alpha = \gamma$. □

Considering the case where $\lambda = i\eta$ yields the following consequence of Theorem 5, which includes Lemma 3 of Wang (1980) as a special case (in the statement of which I think $\geq$ should be $>$).

Corollary 5. *If* $E(\int X(dx))^{\gamma} < \infty$ *and* $m(0) < |m(i\eta)|^{\gamma}$, *then* $\{W^{(t)}(i\eta)\}$ *converges almost surely and in* γ*th mean.*

As indicated above, Theorem 5 requires no further proof; a skeleton argument and standard martingale properties suffice. However, the analogue of Theorem 2, giving uniform convergence, which is now stated, does require a little additional argument.

Theorem 6. $\{W^{(t)}(\lambda)\}$ *converges uniformly on any compact subset of* Λ, *almost surely and in mean, as* $t \to \infty$.

Proof. Let F be a compact subset of Λ; then $W^{(t)}$ is a martingale in the Banach space of continuous functions on F with the supremum norm (denoted by $\|\cdot\|$). Furthermore, by Theorem 2, $W^{(n\delta)}$ has a limit W in this space as $n \to \infty$ through the integers for any $\delta > 0$. It follows that $E\|W^{(t)} - W\| \to 0$ as $t \to \infty$, giving the convergence in mean, and that $\|W^{(t)} - W\| \to 0$ a.s. as

$t \to \infty$ through the rationals. This will imply the stated result if $\|W^{(t)} - W\|$ is regular, which can be verified by routine analysis. □

Consider now the continuous-time analogues of the large deviation results Theorem 4 and Corollary 4, obtained by simply replacing n by t throughout their statements. [This will include a uniform version of Theorem 1 of Uchiyama (1982) as a special case.] The first part of the proof of Lemma 5 needs only obvious changes, with Theorem 6 being invoked in place of Theorem 2. However, the second part does require some work. In continuous time the crucial estimate (4.4) must be refined further if we are to get convergence for all t.

Let $I_n = \{t: n < t \le n + 1\}$. We replace (4.4) by

$$\sup_{\theta \in B_r, t \in I_n} t^{p/2} \int_{\varepsilon \le |\eta| \le a} |W^{(t)}(\lambda)\phi(\lambda)^t| \, d\eta$$

$$\le K(n+1)^{p/2} \left\{ \frac{\sup\{|B^{(t)}(\lambda)|: \lambda \in G_r, t \in I_n\}}{\inf\{m(\theta)^{n+1}: \theta \in B_r\}} + \delta^n \right\}.$$

Refining (4.5) and (4.6) similarly gives

$$E \sup_{\lambda \in G_r, t \in I_n} |B^{(t)}(\lambda)| \le K \sup_{\lambda \in G_{2r}} \left(E \sup_{t \in I_n} |B^{(t)}(\lambda)|^{\alpha} \right)^{1/\alpha}. \tag{5.4}$$

Now observe that, by (5.1), $m(\lambda)$ is never 0, and $|B^{(t)}(\lambda)/m(\lambda)^t|$ is a regular submartingale. Hence, by a standard martingale inequality [Williams (1979), Lemma 43.3],

$$E \sup_{t \in I_n} \left| \frac{B^{(t)}(\lambda)}{m(\lambda)^t} \right|^{\alpha} \le KE \left| \frac{B^{(n+1)}(\lambda)}{m(\lambda)^{n+1}} \right|^{\alpha}$$

and so (5.4) is less than

$$K \sup_{\lambda \in G_{2r}} \left(E |B^{(n+1)}(\lambda)|^{\alpha} \right)^{1/\alpha}.$$

The proof now continues as in the discrete case.

Acknowledgment. I would like to thank an anonymous referee for his most perceptive and useful report. In particular, he drew my attention to the paper by Stone (1967), thereby sparing the reader my own laboured version of a local large deviation theorem and simplifying the task of giving the uniform versions recorded here.

REFERENCES

BIGGINS, J. D. (1977). Martingale convergence in the branching random walk. *J. Appl. Probab.* **14** 25–37.

BIGGINS, J. D. (1978). The asymptotic shape of the branching random walk. *Adv. in Appl. Probab.* **10** 62–84.

BIGGINS, J. D. (1979). Growth rates in the branching random walk. *Z. Wahrsch. Verw. Gebiete* **48** 17–34.

BIGGINS, J. D. (1989). Uniform convergence of martingales in the one-dimensional branching random walk. In *Selected Proceedings of the Symposium on Applied Probability, Sheffield 1989* (I. V. Basawa and R. L. Taylor, eds.) **18** 159–173. IMS, Hayward, Calif.

HÖRMANDER, L. (1973). *An Introduction to Complex Analysis in Several Complex Variables.* North-Holland, Amsterdam.

JOFFE, A., LE CAM, L. and NEVEU, J. (1973). Sur la loi des grandes nombres pour les variables aléatoires de Bernoulli attachées à un arbre dyadique. *C. R. Acad. Sci. Paris Sér. A–B* **277** 963–964.

JOFFE, A. and MONCAYO, A. R. (1973). Random variables, trees and branching random walks. *Adv. in Math.* **10** 401–416.

KINGMAN, J. F. C. (1975). The first birth problem for an age-dependent branching process. *Ann. Probab.* **3** 790–801.

NEVEU, J. (1975). *Discrete Parameter Martingales.* North-Holland, Amsterdam.

NEVEU, J. (1988). Multiplicative martingales for spatial branching processes. In *Seminar on Stochastic Processes, 1987. Progress in Probability and Statistics* (E. Çinlar, K. L. Chung and R. K. Getoor, eds.) **15** 223–241. Birkhäuser, Boston.

STAM, A. J. (1966). On a conjecture of Harris. *Z. Wahrsch. Verw. Gebiete* **5** 202–206.

STONE, C. (1967). On local and ratio limit theorems. *Proc. Fifth Berkeley Symp. Math. Statist. Probab.* **2**(ii) 217–224.

UCHIYAMA, K. (1982). Spatial growth of a branching process of particles living in $\mathscr{R}^d$. *Ann. Probab.* **10** 896–918.

VON BAHR, B. and ESSEEN, C.-G. (1965). Inequalities for the rth absolute moment of a sum of random variables $1 \le r \le 2$. *Ann. Math. Statist.* **36** 299–303.

WANG, F. J. S. (1980). The convergence of a branching brownian motion used as a model describing the spread of an epidemic. *J. Appl. Probab.* **17** 301–312.

WATANABE, S. (1967). Limit theorem for a class of branching processes. In *Markov Processes and Potential Theory* (J. Chover, ed.) 205–232. Wiley, New York.

WILLIAMS, D. (1979). *Diffusions, Markov Processes, and Martingales.* Wiley, New York.

DEPARTMENT OF PROBABILITY AND STATISTICS
UNIVERSITY OF SHEFFIELD
P.O. BOX 597
SHEFFIELD S10 2UN
ENGLAND